

PREDICTION OF FUNCTIONAL EFFICIENCY QUALITY WITH RANDOM FOREST: SOFTWARE INDUSTRY-A CASE STUDY

Sajal Kanti Das¹, Jamal Hussain² and Piya Das³

^{1,2}Department of Mathematics and Computer Science, Mizoram University, India

³Department of Economics, Radhamadhab College, India

Abstract

The software industry in India is a major contributor for the service sectors in the country. Efficiency initiatives in this area are crucial. The present study evaluates a sensitivity function technique for analyzing the performance of twelve software company. A performance evaluation metric called the Integrated Functional Efficiency Quality (IFEQ) is assessed. The performance trends of the selected software company is compared. Random Forest, machine learning technique is employed to forecast the IFEQ for the next five years. The performance of Random Forest in order of evaluation metric is presented. As revealed by experimental results most of the companies are at average level in order of IFEQ. As inference a benchmark company is chosen. The less efficient company can follow the benchmark one. Managerial decision making may be supported in policy design and adoption. Such practice may yield business efficacy.

Keywords:

Performance Analysis, Functional Efficiency Quality, Software Industry, Decision Making, Benchmarking

1. INTRODUCTION

The IT sector comprising the software is one of the major sector India which is solely favored by the available of vast skilled IT professional, cheap labor and favorable govt policies. The global software markets demand determines the ups and down in the Indian IT market. outsourced based service from India, is a support for the clients in developed countries. The Indian entrepreneurs set up software development firms to cater such need by global IT market. Tata Consultancy Services (TCS), Cognizant Technology Solutions, Infosys Technologies, Wipro, HCL Technologies are India software companies which made a global presence and are continuing the legacy yet.

The software industry comprises of the IT services with Application Maintenance and Support which mainly provide services to maintain and support existing software applications for clients worldwide and also Indian companies excel in providing system integration service software industry also provides various engineering services like product design, software prototyping, and software testing for clients in several sectors like automotive, aerospace, and telecommunications.

The speedy delivery, internationalization and adapting the pace with evolving technology are the facts behind the growth of software industry, Jain *et al.* [9]. There are various studies being taken for the performance assessment of the software industry. These attempts include the efficiency analysis with the frontier and non-frontier-based methods such as by employing data envelopment analysis, total factor productivity etc. The present experiment attempts to assess the integrated functional efficiency quality of some selected software firms in India. Rest of the paper proceeds as follows. Section 2 elaborates brief literature in

software firm's efficiency. Section 3 describes some underlying theoretical concepts, followed by section 4 which presents the results (IFEQ assessment and prediction) and section 5 concludes the paper.

2. LITERATURE REVIEW

The software industry has experienced rapid expansion and constitutes a critical component of India's service-driven economy, contributing substantially to gross domestic product (GDP), foreign exchange earnings, and large-scale employment generation. Given its strategic economic significance, the assessment of efficiency within this sector assumes considerable importance, particularly in understanding how effectively resources are utilized to generate output and sustain competitive advantage.

A number of empirical studies have examined the efficiency dynamics of the Indian software industry using diverse methodological frameworks. For instance, Mathur [2] employed an input-oriented Data Envelopment Analysis (DEA) model to evaluate the technical efficiency of 92 software firms in India during the period 2005–2006, with a particular focus on identifying the determinants influencing efficiency variations across firms. Similarly, Bhattacharjee [3] analysed the technical efficiency of firms operating within the Software Technology Parks (STPs) in Kolkata by adopting an output-oriented DEA model under the Variable Returns to Scale (VRS) assumption, thereby capturing scale-related efficiency differentials.

Further contributions include Sahoo [4], who estimated Total Factor Productivity (TFP) growth for 72 selected software companies over the period 1998–2008 using the Malmquist Productivity Index. The findings indicated a modest TFP growth of approximately 0.4 percent during the study period. Extending this line of inquiry, Sahoo and Nauriyal [5] examined patterns of technical efficiency in Indian software firms from 1999 to 2008 through an input-oriented DEA model under VRS assumptions, revealing that market structure and competitive intensity exert significant influence on firm-level efficiency outcomes.

Subsequent studies have adopted more advanced analytical frameworks to deepen the understanding of efficiency dynamics. Das [6] and Das and Datta [7] utilized a two-stage DEA approach to investigate both the trends and determinants of technical efficiency in the Indian IT and IT-enabled services (ITeS) industry over the period 2000–2014. These studies incorporated Pareto–Koopmans efficiency measures alongside CCR and BCC models, as introduced by [1], to account for input and output slacks more comprehensively. In a related context, [8] assessed the role of policy interventions, particularly tax incentives and subsidies, and found that such measures positively influence the efficiency of the software industry. Furthermore, [9] reported a

relatively high level of TFP within the Indian software sector, indicating efficient utilization of inputs in output generation.

More recently, Goyal et al. [10] evaluated the relative efficiency of 18 Indian IT service firms using a DEA framework, wherein labor effort was considered as the sole input variable. Their findings underscore the critical role of human capital in driving efficiency within knowledge-intensive service industries. Collectively, these studies highlight the evolving methodological landscape and underscore the multifaceted determinants of efficiency in the Indian software sector.

3. THEORETICAL CONCEPTS

3.1 EFFICIENCY

For the optimal uses of resources or in the other word in the process of input to output conversion to assess the quality of the process of conversion a metric is considered which can be efficiency. Efficiency is defined as the ratio of total output to total input. The cost of the product with respect to the total cost involved to produce the particular product. Efficiency can be represented as in Eq.(1).

$$\text{Efficiency} = \text{Output/Input} \quad (1)$$

3.2 FUNCTIONAL EFFICIENCY QUALITY (FEQ)

The idea of functional efficiency quality is based on the relative comparison of the parameters value exhibited by the companies. Functional efficiency can be defined as the overall functioning of the company taking into account all of the variables present in the data set. The overall efficiency of the company's business will be assessed with the functional efficiency quality considering all perspectives.

The value of a parameter for a company is compared to the best in case of the parameter which is favorable for the efficiency measure. The notion of functional efficiency quality (FEQ) proposed by Sajal and Hussain [12] to assess the of businesses within an industry is called functional efficiency quality. IFEQ is calculated here for all the year. All the variable presents the dataset contributes to the FEQ measure. this Hence the FEQ measeure is data driven. For evaluating the functional efficiency quality sensitivity function defined in Sundararajan [1] is used. Sensitivity numbers are the values of the sensitivity function that are acquired for each of the company's selected parameters. With a mapping $fp: [-1, +1]$, these sensitivity number values can be transferred to the real domain R .

3.3 INTEGRATED FUNCTIONAL EFFICIENCY QUALITY (IFEQ)

The Integrated Functional Efficiency Quality (IFEQ) is an integrated metric that depicts the overall functioning of a company. It is estimate considering each parameter in the dataset. In a given year for a considered company IFEQ incorporates the overall operating or efficiency quality of that business. It is the total value of a company's quality index for a specific year that has been chosen. The algorithm for assessing is proposed in Sajal and Hussain [12].

There are some sequences of steps followed to calculate the Functional efficiency quality. The number called quality index is

assessed for calculating the sensitivity number using the sensitivity function. These quality index which is calculated for each parameter present in the dataset is aggregated to compute the IFEQ. Each parameter can be marked favorable and non-favorable on the basis of their value. The algorithm for IFEQ assessment is detailed in Sajal and Hussain [12]. The assessment procedure includes two separate formulae for the favorable and unfavourable parameters. Initially the parameter from the dataset are identified as favorable and non-favorable. The quality index of each parameter is then assed using the sensitivity function. Eventually IFEQ is obtained by integrating the quality index of each parameter.

4. IFEQ ASSESSMENT AND PREDICTION

For the assessment of IFEQ twelve software company are selected are presented in Table.1. the mean IFEQ values for all these selected companies over thirty years is provided in figure 1. These estimated IFEQ from year 1992 to 2021 are used to train the model for random forest. IFEQ is estimated and considered for training the Random Forest machine learning model. The Random Forest is a machine learning ensemble method that can be used for regression and classification task. The random forest combines various individual decision trees to give a robust and accurate prediction. The building block of a random forest model is decision tree. The Bootstrap Aggregating known as Bagging which involves making several subsets form the original training data by employing random sampling with replacement. Random forest machine learning technique is used to estimate the IFEQ. The traditional IFEQ values are used as a basis to predict the IFEQ for the unknown years i.e. from year 2022 to 2026.

Table.1. Selected companies

Sl. No.	Company Name
1	Capricorn Systems Global Solutions Ltd.
2	Digispice Technologies Ltd.
3	E I T Services India Pvt. Ltd.
4	Infosys Ltd.
5	Mastek Ltd.
6	NIIT Ltd.
7	R S Software (India) Ltd.
8	Silverline Technologies Ltd.
9	Tata Elxsi Ltd.
10	Tech Mahindra Ltd.
11	Wipro Ltd.
12	Zensar Technologies Ltd.

The Fig.1 presents the comparison of traditional IFEQ with the IFEQ computed with random forest ensemble method. The Random Forest model closely reproduces the observed mean IFEQ values for all companies, indicating a high degree of consistency between the estimated and original efficiency-quality measures. Similarly Fig.2 compares the actual IFEQ with predicted IFEQ with random forest for all companies for the study period. The scatter plot shows a strong agreement between the actual and predicted IFEQ values, with most observations closely

aligned along the ideal prediction line. This indicates that the Random Forest model accurately captures the underlying relationship in the dataset and produces reliable IFEQ estimates across different companies and years. Although a few observations deviate slightly from the reference line, the overall dispersion is minimal, suggesting low prediction error and good predictive performance. It demonstrates the effectiveness of the Random Forest algorithm for IFEQ estimation in the Indian software industry. The throughout assessment of all companies for the study period is presented in Appendix A at the end of the paper. It is observed that for Capricorn Systems Global Solutions Ltd. for the IFEQ remained stable historically and the same is predicted by the Random Forest model during 2022-2026. Differently Digispice Technologies Ltd. witnessed fluctuation, but forecasting indicates stable functional efficiency trend over the prediction period. For E I T Services India Pvt. Ltd. efficiency got improved in the historical phase and the prediction indicates uniform stable performance with positive trends.

- In case of Infosys Ltd., a significant recovery in historical IFEQ, while model predicts stable efficiency from 2022–2026. IFEQ was at peak from 2010-2012 for Mastek Ltd. and the forecast indicates a stable but slightly lower efficiency level. For NIIT Ltd. IFEQ is table in 2011 and forecasted IFEQ is constant from 2022-2026. The functional efficiency of R S Software (India) Ltd. increased largely in 2014, followed by a gradual decline and prediction shows a minor downward trend.
- Silverline Technologies Ltd. maintained a stable IFEQ throughout the study period, and the Random Forest model predicts consistent functional efficiency during 2022–2026 with only marginal variation.
- Tata Elxsi Ltd. exhibited an overall improvement in historical IFEQ, while the forecast indicates a stable functional efficiency trend over the prediction period.
- Tech Mahindra Ltd. demonstrated consistently high IFEQ during the historical period, and the Random Forest model forecasts sustained functional efficiency during 2022–2026.
- Wipro Ltd. recorded the highest historical IFEQ among the selected companies, with the forecast indicating continued high functional efficiency and minimal fluctuations.
- Zensar Technologies Ltd. showed a gradual improvement in IFEQ over the historical period, and the Random Forest forecast suggests stable functional efficiency throughout 2022–2026.

Table.2. Model summary of random forest

Model	MSE	MSE	MSE	R ²	R ²	R ²
	mean	std	max	mean	std	max
Random forest	0.468	0.548	0.948	0.650	0.540	0.974

The IFEQ assessment with RF and without RF reveals almost same measure. While considering the IFEQ predicted with random forest for the year 2022 to 2026, Wipro Ltd. with 27.15 is having the highest score, followed by Tech Mahindra Ltd. with 24.28, EIT Services India Pvt. Ltd. having 23.45 is at third best performer. Digispice Technologies Ltd. with 21.04 is at bottom, preceded by NIIT Ltd., Silverline Technologies Ltd., and Capricorn Systems Global Solutions Ltd. The four average companies are Infosys Ltd., R S Software (India) Ltd., Zensar

Technologies Ltd. and Mastek Ltd. The mean IFEQ predicted for twelve company is 22.48. Table 2 presents the model summary of random forest which represent the Mean Square Error (MSE), Adjusted R-squared denoted as R2.

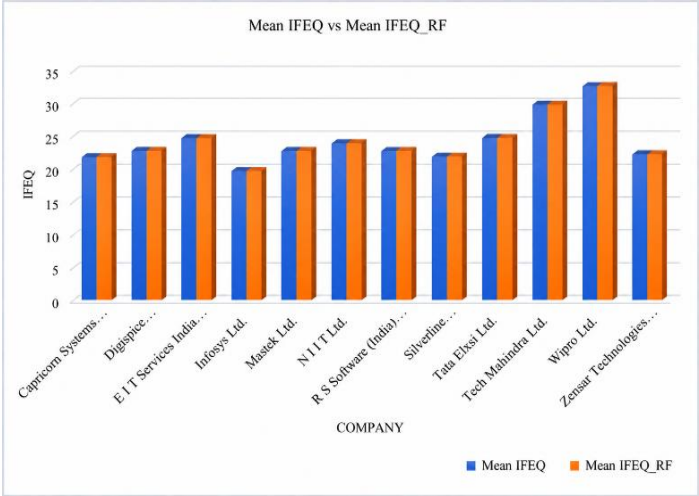


Fig.1. Mean IFEQ compared with mean IFEQ with Random Forest for all Companies from 1992 to 2021

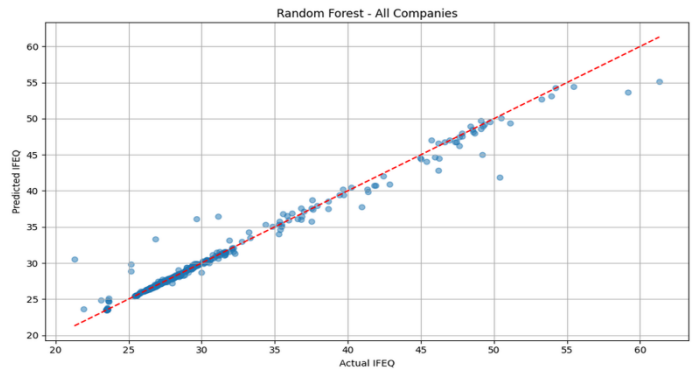
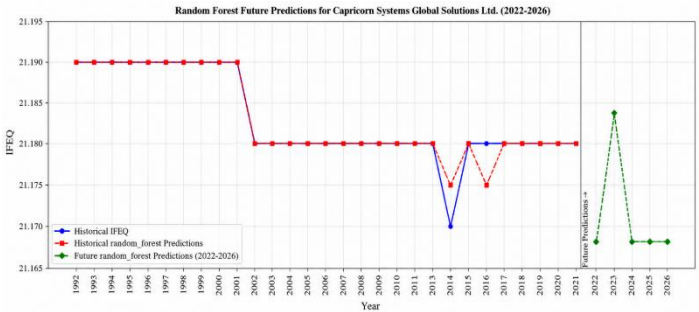
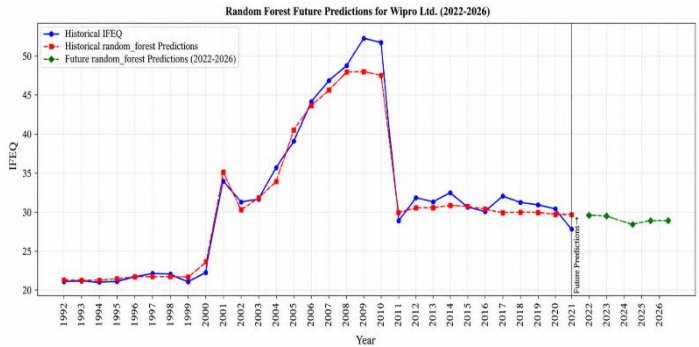
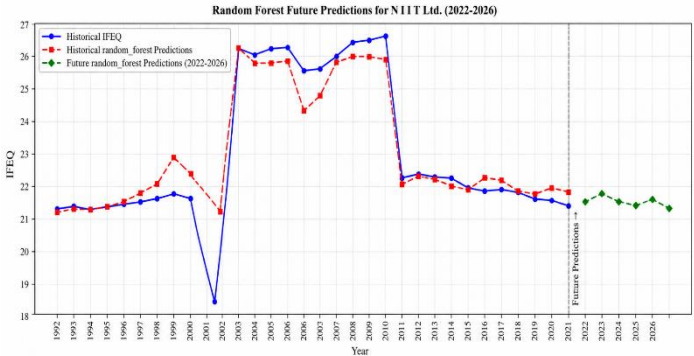
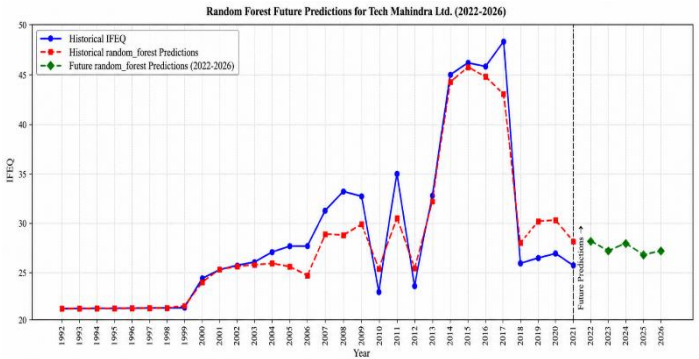
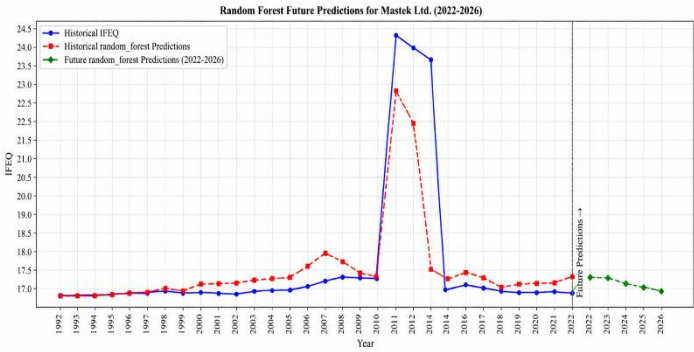
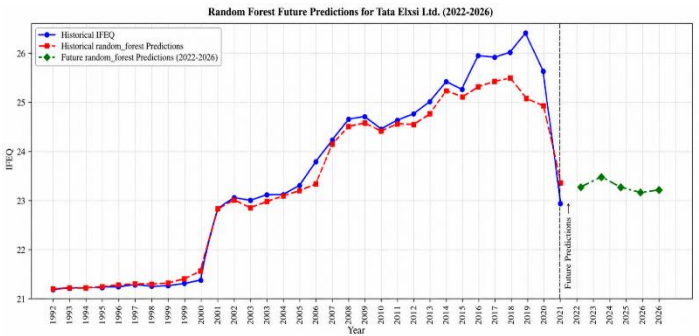
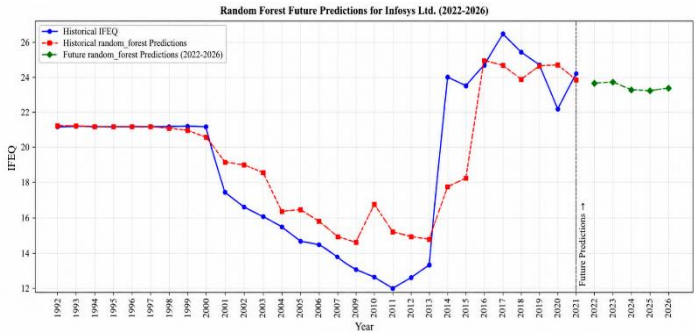
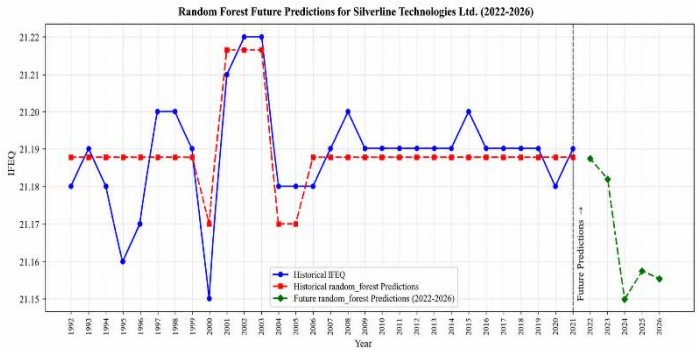
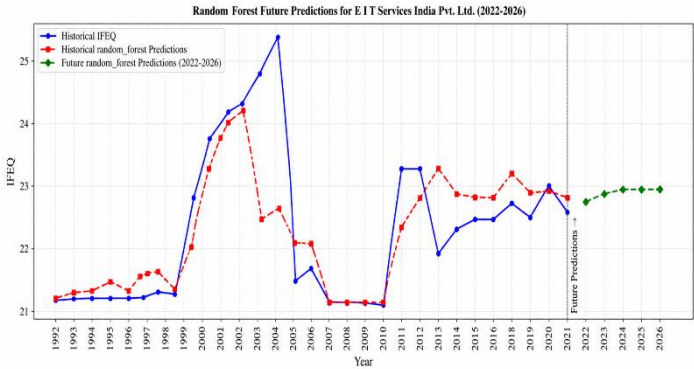
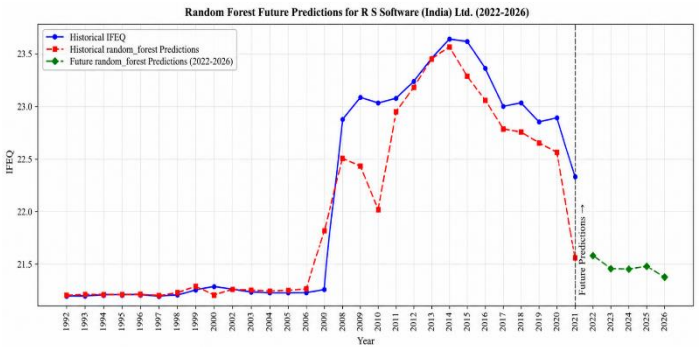
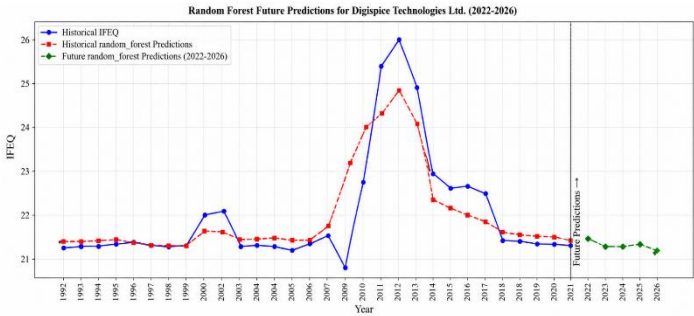


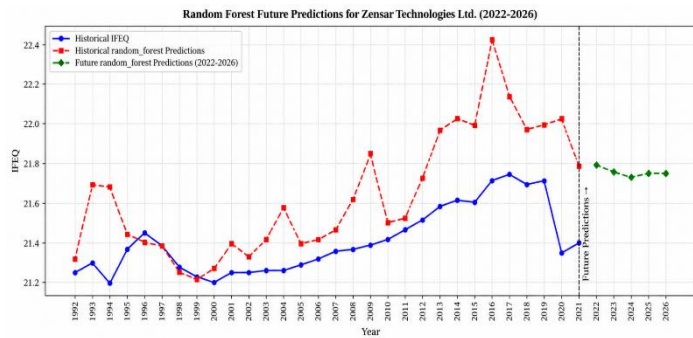
Fig.2. Actual IFEQ compared with IFEQ Random Forest for all Companies from 1992 to 2021

Appendix A

The detailed plot of the historical IFEQ, IFEQ with random forest and the predicted IFEQ from 2022-2026 for all 12 companies are presented below:







5. CONCLUSION

From the underlying analysis when the mean IFEQ is considered for each company for the 30 years it is observed that with IFEQ 31.69 Wipro Ltd. comes at the top followed by Tech Mahindra Ltd. with 28.23. Wipro Ltd., Tech Mahindra Ltd., E I T Services India Pvt. Ltd. and Tata Elxsi Ltd. are the for best performing company. Infosys Ltd. with 19.09 came out as the most inefficient company over the considered thirty years. Infosys Ltd., Zensar Technologies Ltd., Silverline Technologies Ltd. and Capricorn Systems Global Solutions Ltd. are bottom four companies. Other four companies are on average performance in order of IFEQ measure. The same result is obtained for IFEQ estimated with RF machine learning techniques.

REFERENCES

- [1] M. Sundararajan and V.J. Loveson, "An Approach Towards the Development of an Environmental Quality Index for Evaluation and Categorization of Environmental Impacts", *Journal of Scientific and Industrial Research*, Vol. 61, pp. 219-223, 2002.
- [2] S.K. Mathur, "A Performance Analysis using Data Envelopment Analysis and Malmquist Index", *Global Economy Journal*, Vol. 7, No. 2, pp. 1-12, 2007.
- [3] S. Bhattacharjee, "Efficiency Dynamics and Sustainability of the Indian IT-ITeS Industry: An Empirical Investigation using DEA", *IIMB Management Review*, Vol. 24, No. 4, pp. 203-214, 2012.
- [4] B.K. Sahoo, "Total Factor Productivity of the Software Industry in India", Available at <https://iegingia.org/upload/annual-report/ar201314.pdf>, Accessed in 2013.
- [5] B.K. Sahoo and D.K. Nauriyal, "Trends in and Determinants of Technical Efficiency of Software Companies in India", *Journal of Policy Modeling*, Vol. 36, pp. 539-561, 2022.
- [6] P. Das, "An Assessment of Performance of Indian Software Industry during 2000-01 to 2014-15 using Data Envelopment Analysis", *International Journal of Engineering, Applied and Management Sciences Paradigms*, Vol. 44, No. 1, pp. 7-21, 2017.
- [7] P. Das and A. Datta, "Performance Evaluation of Indian Information Technology-Enabled Services (ITeS) Industry: An Application of Two-Stage Data Envelopment Analysis", *International Journal of Advances in Management and Economics*, Vol. 6, No. 2, pp. 52-70, 2017.
- [8] Das and Prosenjit, "An Evaluation of the Determinants of Total Factor Productivity Growth in Indian Information Technology Industry: An Application of DEA-based Malmquist Index", *The Central European Review of Economics and Management (CEREM)*, Vol. 1, No. 4, pp. 175-224, 2017.
- [9] N.K. Jain, S. Celo and V. Kumar, "Internationalization Speed, Resources and Performance: Evidence from Indian Software Industry", *Journal of Business Research*, Vol. 95, pp. 26-37, 2019.
- [10] S. Goyal, A.N. Sah, R.K. Sharma and J. Puri, "Estimating Technical Efficiencies of Indian IT Companies for Setting Improvement Targets for Inefficient Companies: An Empirical Analysis with Workers' Effort as Key Input", *Work*, Vol. 66, No. 4, pp. 885-900, 2020.
- [11] "Economic Survey 2023-24", Government of India, Available at: <https://www.indiabudget.gov.in/economicsurvey/doc/echapter.pdf>, Accessed in 2024.
- [12] Sajal Kanti Das and Jamal Hussain, "Functional Efficiency Quality of Indian Software Industry: A Computational Approach", *Journal of Science and Technology*, Vol. 12, No. 2, pp. 68-72, 2024.